Nucleic Acids Research, 2005, Vol. 33, No. 11 ***e99***
doi:10.1093/nar/gni098

High-throughput alternative splicing quantification by primer extension and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

Ron M. McCullough, Charles R. Cantor[1] and Chunming Ding[2,*]

Program of Molecular and Cellular Biology and Biochemistry, [1]Center for Advanced Biotechnology, 36 Cummington Street, Boston, MA 02215, USA and [2]Centre for Emerging Infectious Diseases, 2/F, School of Public Health, Faculty of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong Special Administrative Region

Received April 10, 2005; Revised May 20, 2005; Accepted June 2, 2005

ABSTRACT

Alternative splicing is a significant contributor to transcriptome diversity, and a high-throughput experimental method to quantitatively assess predictions from expressed sequence tag and microarray analyses may help to answer questions about the extent and functional significance of these variants. Here, we describe a method for high-throughput analysis of known or suspected alternative splicing variants (ASVs) using PCR, primer extension and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). Reverse-transcribed mRNA is PCR amplified with primers surrounding the site of alternative splicing, followed by a primer extension reaction designed to target sequence disparities between two or more variants. These primer extension products are assayed on a MALDI-TOF mass spectrometer and analyzed automatically. This method is high-throughput, highly accurate and reproducible, allowing for the verification of the existence of splicing variants in a variety of samples. An example given also demonstrates how this method can eliminate potential pitfalls from ordinary gel electrophoretic analysis of splicing variants where heteroduplexes formed from different variants can produce erroneous results. The new method can be used to create alternative variant profiles for cancer markers, to study complex splicing regulation, or to screen potential splicing therapies.

INTRODUCTION

Since the discovery of alternative splicing in 1977 (1), our estimate of the degree to which pre-mRNA undergoes alternative splicing increases every year as more studies are conducted. In 1994, alternative splicing was thought to affect a mere 5% of genes in eukaryotes (2), five years later the number jumped to 35% (3), and now, 28 years after its discovery some researchers using microarrays have estimated that 74% of human multi-exon genes are alternatively spliced (4) with 80% of these alternative splicing events producing changes in protein sequences (5). Regardless of the final tally, the importance of alternative splicing is assured; for every role of alternative splicing that we know of, there are many more being uncovered. Examples of the functions of alternative splicing include calcium pump diversity (6) and regulation of fibronectin (7), calcitonin (8) and GTP cyclohydrolase I (9). In addition to its intended role in complex regulation of signaling and expression, alternative splicing has been linked to many well-characterized diseases, including cystic fibrosis, Alzheimer disease (10), Parkinson's and Frasier syndrome (11) as well as every major cancer (12). How these variants are regulated through *cis*-regulatory elements and *trans*-acting factors and how this knowledge can be used in diagnosis, prognosis and treatment of the aforementioned diseases is an active field of study. With such vast and significant roles linked to alternative splicing, the need to fully characterize alternative splicing qualitatively and quantitatively is imperative.

To understand the impact of alternative splicing on the complexity of the human genome, researchers have used expressed sequence tag (EST) analysis and microarray profiling to determine possible variants that are tissue-specific,

*To whom correspondence should be addressed. Tel: + 852 2252 8842; Fax: + 852 2635 4977; Email: cmding@cuhk.edu.hk

© The Author 2005. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oupjournals.org

tumor-specific and disease-associated (12–16). These searches have identified thousands of potential tissue-specific variants and cancer-specific markers. Unfortunately, current methods for verifying and quantifying these variants at the level needed are either too low throughput or not sufficient in quantitative accuracy. Among many technologies that have been used for alternative splicing analysis, polymerase colony technology (17) was shown to be a particularly powerful method to analyze extremely complex alternatively spliced genes, such as CD44 or MAPT. However, it is unclear how the polony technology can be adapted to clinical situations where high-throughput analysis of hundreds of potential targets, scores of tissue types and hundreds to thousands of samples are needed to generate statistical significance without being labor intensive and cost prohibitive.

Here, we demonstrate one approach for alternative splicing variant (ASV) quantification. Complementary DNA (cDNA) reverse transcribed from total RNA preparation is PCR amplified using primers that flank the site of alternative splicing. To distinguish different ASVs, an extension primer is designed to extend from one to a few bases into the sequence differences introduced by alternative splicing, such that the different variants create products of different molecular weights. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is then used to detect and quantify these products. The entire process is carried out in a 384-well plate format in a highly automated fashion so that a single person can carry out a few thousand reactions per day.

MATERIALS AND METHODS

Samples and reverse transcription

Total RNA samples from six healthy human tissue samples, including brain, colon, heart, kidney, liver and testis, were obtained from Clontech. A sample of 100 ng of RNA was reverse transcribed for 1 h at 42°C with 0.5 μg of random hexanucleotides and an AMV reverse transcriptase (Promega) in 20 μl total volume.

PCR primer design, amplification and primer extension

Primers were designed using Sequenom's Assay Designer software (Table 1) and obtained from Integrated DNA Technology (Coralville, IA). PCR primers were tagged with 5′-ACGTTGGATG-3′ at the 5′ end to avoid interference with the mass spectra (tag sequence is not shown in Table 1). Amplification of 10 ng of cDNA was performed using PCR primers at 100 nM, $MgCl_2$ at 2.75 mM and 200 μM dNTP using 0.1 U HotStart *Taq* DNA polymerase (Qiagen) in 5 μl with the following PCR conditions: 95°C hot start for 15 min, followed by 45 cycles of 95°C for 30 s, 56°C for 1 min, then 72°C for 1:30 min, with a final hold of 72°C of 7 min. After the PCR amplification, the products were treated with 0.04 U shrimp alkaline phosphatase, SAP (Sequenom), which inactivates unused dNTPs from the amplification cycles, for 20 min at 37°C followed by heat inactivation at 85°C for 5 min. For the extension cycle, 1.2 μM final concentration of extension primer and 0.6 U of ThermoSequenase (Sequenom) were added to a total reaction of 9 μl with the termination mixture containing specific dideoxynucleotides and deoxynucleotides for each reaction at 50 μM for each base. The extension conditions include a 94°C hold for 2 min with 75 cycles of the following: 94°C for 5 s, 52°C for 5 s and 72°C for 5 s.

MALDI-TOF MS and quantitative analysis

Prior to MALDI-TOF MS analysis, salts from the reactions were removed using SpectroCLEAN resin and 16 μl of water. ASV analysis was performed using the MassARRAY system (Sequenom) by dispensing ≈10 nl of final product onto a 384-plate format MALDI-TOF MS SpectroCHIP using a SpectroPOINT nanodispenser (Sequenom). The frequency of each variant was generated by SpectroTYPER (Sequenom) and was determined by taking the peak area ratio of the peak associated with the splice variant of interest over the

Table 1. Primer and termination mixes for ASV analysis

Gene[a]	PCR primer 1	PCR primer 2	Extension primer	Termination mixture[b]
HNRPA2B1, NM_031243	AGATTCTCTCTCATCTCGCTC	TCCCTCATTACCACACAGTC	GAACTGTTCCTTTTCTCTCT	CG
ACTN-4, NM_004924	ACAGACCAGAGCTGATTGAG	CTGTCAACCAAGAGAACGAGCA	ATGCTGGATGCAGAGGA	ACG
BCL2L1, NM_001191	ATCAATGGCAACCCATCCTG	GGTCAGTGTCTGGTCATTTC	TCAGAGCTTTGAACAGG	CT
Casp9, NM_001229	AGACCAGAGATTCGCAAAC	TTCCTGGAACGGGGTGGCAT	AGGATTTGGTGATGTCG	AT
Hells, AB102716	AAAAGGCTCGCATGTCTTGG	ACTGGCTTCTCTTCACTTGC	TGGAACAGCAACAATTAGAG	AT
Keratin 8, NM_002273	TGAACAAGGTAGAGCTGGAG	GCTGAGAGCATGTACCAGAT	GTCCCAGATCTCGGACA	CT
NME1, NM_000269	GCAAGACAGGTTCACAATGG	CAAGCTTCCGAAGATCTTCT	AACCTAAGCAGCTGGAAGG	AGT
PCBP4, NM_020418	CTCACAGGAGTTCTTGGTTC	CATTAGCTGCTGCCATCTTG	CCAGTACCTCATCACTGCCTG	CT
PPP1R11, NM_170781	GAGACCGTCACTGAGACAAC	AAAGGCCCGAGGTTTCTCAT	CCGTGACAACCGAGCCCGA	ACG
RAB1A, NM_004160	ACAGAAAGCTACATCAGCAC	GACGATGGCAGCTGAGATTA	GAAAACAATCAAGCTTCAAATA	ACT
REST, AF228045	GAACTCATACAGGAGAACGC	GCAAGACAGGTTCACAATGG	ACATATGCGTACTCATTCAG	CT
TERT				
Alpha NM_198255	CGCCTGAGCTGTACTTTGTC	TGAGGTGAGGTGTCACCAAC	GACCTCCGTGAGCCTGTCCT	CG
Beta NM_196235	CGCCTGAGCTGTACTTTGTC	TGAGGTGAGGTGTCACCAAC	CCTTCAAGAGCCACGTC	CG
VEGF, NM_003376				
Set one	AAACCCTGAGGGAGGCTCCT	TGCTGTCTAATGCCCTGGAG	GCCTCGGCTTGTCACAT	CG
Set two	CGAGTCTGTGTTTTTGCAGG	GGCAAAAACGAAAGCGCAAG	GAGCAAGGCCCACAGGGA	AGT

Accession numbers are for the longest splice variant mRNA, except TERT where the accession numbers are for the specific isoforms.
[a]Sequences for cDNA obtained from BLAT database (http://genome.ucsc.edu/).
[b]Example termination mixture CT indicates dATP, ddCTP, dGTP and ddTTP.

total area (ASV area + normal variant area). Furthermore, the data quality was controlled by discarding data with a frequency error (a weighted uncertainty of the frequencies for each variant reported by the SpectroTYPER software) >10.0%. Level of quantification (LOQ) for the MALDI-TOF platform has been previously reported using pooled DNA as 5.0% (18). Although some of the assays investigated here demonstrated reproducible data at lower frequencies, the conservative value of 5.0% was used as the LOQ. Although this method can detect ASVs <5%, the detection is not quantitative at this stage. A peak was considered detected if it had a signal-to-noise ratio >10. After the data were filtered for quality, the mean and standard deviation for each splice variant was calculated as the final reported value.

ACTN-4 analysis

The ACTN-4 PCR product and the 1:20 dilution samples were analyzed using a 10% TBE non-denaturing pre-cast acrylamide gel (Invitrogen) using 0.5× TBE buffer. The ACTN-4 PCR product was produced by using the PCR amplification step outlined in the previous assay using 100 ng of tissue cDNA sample in a total reaction of 20 µl. The 1:20 samples were obtained by adding 1 µl of the ACTN-4 PCR product to a total PCR volume of 20 µl followed by a single PCR cycle. Bands were electro-eluted using GeBA flex-tubes according to instructions, and the sequence of each band was performed by Davis Sequencing (Davis, CA) using the same forward and reverse primers used in the PCR.

RESULTS

To evaluate the potential use of the primer extension method for the detection and quantification of ASVs, examples were selected consisting of genes whose variants are well characterized and have been shown to be potential cancer markers (BCL2L1, TERT, VEGF, HNRPA2B1 and CEACAM), a few less-characterized ASVs, (ACTN-4, KRT8, PCBP4, Caspase 9 and HELLS) as well as some ASVs identified by data mining techniques (REST, RAB1A, PPP1R11 and NME1). Although there are many types of ASVs, they all share the fact that a sequence difference is introduced into the mRNA. It is this fact that allowed for the detection and quantification of the splicing events, regardless of the exact nature of the ASV. Throughout this paper, the term 'alternatively spliced variant' will refer to the cancer-associated variant and 'normal' will refer to non-cancer-associated variant. Furthermore, because this analysis uses relative percents the sum of the ASV and normal variant is 100%, thus a 10.4% ASV has 89.6% normal variant present.

For sequence insertions (Figure 1a, c and d), the variant sequence differs from the native sequence only in the novel insertion whether it is a unique series of exons (REST), a 3′ addition, (TERT alpha variant) or a 5′ addition (BCL2L1-L). For sequence deletions (Figure 1b), the ASV is detected in the sequence disparity of the exon adjacent to the fixed exon. Normal splicing for TERT results in extension from exon 6 into exon 7 while the beta variant results in extension into exon 9. For mutually exclusive exons (Figure 1e), as in ACTN-4, the extension primer is designed for the sequence portion of the 8′ exon that is unique from exon 8. Complex splicing, such as vascular endothelial growth factor (VEGF), where multiple splicing events can result in many variants, extension strategies are less straightforward and will be discussed in detail later on.

The different sequences that are produced from the extension primer of each variant were detected as a specific peak in the mass spectra. For example, in BCL2L1-L (also known as BCL-X_L), a CT dideoxy termination mixture (dATP, ddCTP, dGTP and ddTTP) and an extension primer create a specific concatenation of the primer, TCAGAGCTTTGAACAGG (5234.4 *m/z*) plus a terminal ddT (5522.6 *m/z*), while the BCL2L1-S variant (also known as BCL-x_S) extends an AddT on to the extension primer (5835.8 *m/z*). The BCL2L1-L and BCL2L1-S variants are uniquely identified by mass/charge ratios and were detected and quantified.

Detection of ASVs

To test the ability of this method to detect splice variants in 13 selected genes (Table 2), total RNA samples from six different healthy tissues were obtained from Clontech and reverse transcribed into cDNA samples. The resultant cDNAs were PCR amplified using forward and reverse primers that were designed to flank the ASV of interest. After amplification, excess dNTPs were removed using SAP, a crucial step that allows the termination mixture used in the extension cycle to have the desired dideoxynucleotide termination sequence. In the primer extension reaction, a primer was designed that was complementary to the sequence directly adjacent to the sequence difference introduced by an ASV. In the extension cycle, ThermoSequenase was used to add the dNTPs until a termination base was added to create a unique extension sequence. Each sample was assayed in replicates of four with each replicate sampled five times by the mass spectrometer.

The brain had the largest relative percent of ASVs expressed, with 48.4%, followed by testis and heart at 35.5 and 35.7%, respectively. We should caution the readers that our data are mainly a proof of principle to demonstrate the methodology rather than to answer the biological question of alternative splicing in a tissue-specific manner. The general trend of the data in Table 2, however, does correlate with previous EST and tissue-specific alternative splicing studies (13,16), in which brain was found to have the highest level of alternative splicing followed by the testis and heart.

In three of the genes that were assayed, TERT, BCL2L1 and RAB1A, the dominant variant in at least two tissues was the alternatively spliced variant. The most dominantly expressed variant was the beta variant of TERT. TERT is a telomerase reverse transcriptase that is primarily regulated transcriptionally (19), but has also been shown to be regulated through alterative splicing of transcripts (20) and is generally inactive in adult tissues with the exception of male germ cells (21). Only the full-length variant is active as both the alpha and beta variants disrupt the required RT motifs (22). Currently, this assay is unable to identify if both the alpha and beta deletion occur on the same transcript; however, we can extrapolate the presence of normal full-length variant by looking at the values from both assays together. In the heart, 87.5% of the TERT transcript has the beta deletion and no TERT transcripts with the alpha deletion were detected, as a result the remaining 12.5% of the TERT transcript has neither the beta nor the alpha deletions, leaving the possibility that those transcripts

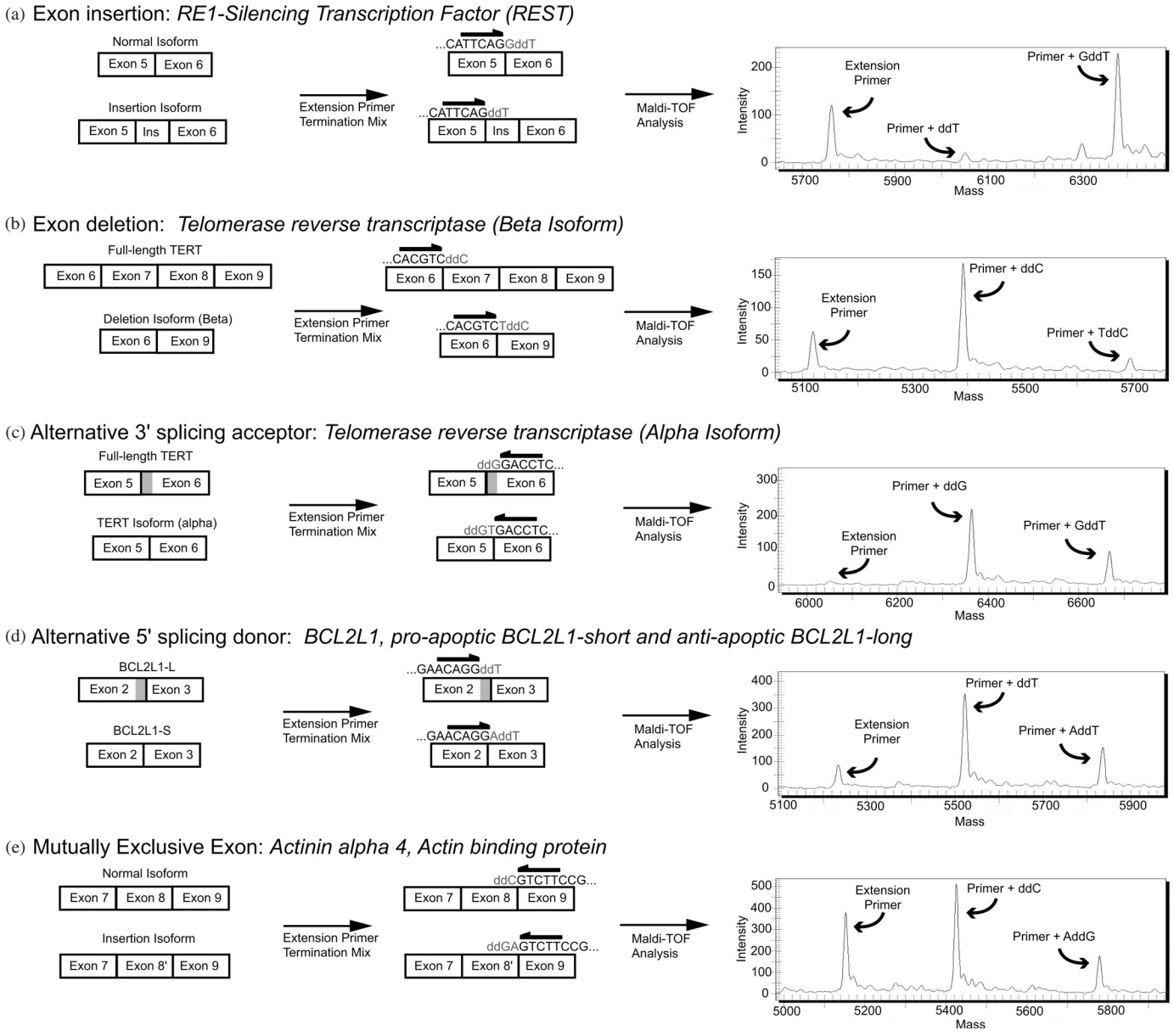

Figure 1. Alternative splicing isoform analysis via primer extension. All classifications of splicing can be analyzed including exon insertion (**a**), exon deletion (**b**), 3′ splicing acceptor (**c**), 5′ splicing donor (**d**) and mutually exclusive exons (**e**). Exon lengths are not to scale. Primer extension assays of respective genes were assayed on a MALDI-TOF MS. In the mass spectra, extension primer refers to unextended primers, while the subsequent peaks indicate the primer with an extension of the base(s) specified.

are full-length. In contrast, >95.0% of the transcripts in the testis tissue contained the beta deletion and 19.9% of those transcripts possessed the alpha deletion, indicating that both the alpha and beta deletion were present on ∼15–20% of the TERT mRNA expressed in the testis. With the exception of the heart, no full-length TERT transcripts were detected in any of the tissues.

BCL2L1-L, a well-known anti-apoptic variant in the Bcl-2 family of cell death regulators, has been linked to a variety of cancers, including lung (23), myeloma (24), brain (25) and breast (26). In most studies, the up-regulation or down-regulation of the BCL2L1-L variant, rather than BCL2L1-S, is monitored. For those studies that do monitor both variants, the values are represented as a ratio of the two variants. Given how well studied this variant is, it is surprising that an alternative splice variant profile in normal tissues has not been presented in previous studies; however, it is clear that the BCL2L1-L is the dominant variant in murine tissue (27). Our data show that of the two variants, BCL2L1-L and BCL2L1-S, BCL2L1-L is close to or above 50% in all samples with the highest percent in the heart with 69.8%.

RAB1A, a member of the RAS oncogene family, ASV tissue profile has not been studied outside of EST computational analysis (28), but our data demonstrate that this alternate variant expressed at near 50% in all tissues with a high of 56.6% in the brain and a low of 40.0% in the testis.

In contrast to the ASVs that were predominantly expressed, four of the genes show either minor ASV expression in all

Table 2. Relative splice variant expression profile in six tissue samples

Gene	Variant (description)	Percent of described splice variant (standard deviation)					
		Brain	Colon	Heart	Kidney	Liver	Testis
HNRP A2B1	B1 (insertion of 36 nt in B1 between exons 1 and 2)	7.1 (1.0)	6.7 (1.4)	6.8 (1.0)	8.2 (1.1)	8.6 (1.9)	6.9 (3.1)
ACTN-4	(Alternative exon 8′ replaces existing exon 8)	71.1 (4.0)	25.7 (2.3)	16.9 (5.2)	32.7 (5.1)	11.9 (3.8)	51.2 (3.8)
BCL2L1	BCL2L1-L (insertion of 189 nt on the 5′ end of exon 2)	61.0 (13.9)	56.0 (9.7)	69.8 (9.8)	65.6 (11.5)	49.1 (11.4)	49.0 (9.7)
CASP9	Caspase 9b (deletion of 450 nt, including exons 3–6)	8.9 (7.2)	<5.0	10.2 (7.2)	8.4 (6.7)	9.4 (7.4)	<5.0
Hells	(Insertion of 44 nt between exons 3 and 4)	<5.0	13.9 (5.3)	<5.0	<5.0	<5.0	<5.0
Keratin 8	(Deletion of 44 nt within exon 5)	ND	7.8 (1.0)	5.3 (0.6)	7.6 (0.4)	7.7 (0.7)	<5.0
NME1	(Insertion of 220 nt between exons 1 and 2)	ND	ND	ND	ND	ND	ND
PCBP4	PCBP4a (deletion of 197 nt from the 3′ end of exon 13)	39.2 (11.4)	21.8 (8.9)	36.2 (10.3)	10.6 (3.8)	13.0 (5.7)	14.9 (6.0)
PPP1R11	(Insertion of 75 nt between exons 1 and 2)	ND	ND	ND	ND	ND	<5.0
RAB1A	(Deletion of 228 nt, including exons 4 and 5)	56.6 (4.7)	45.7 (2.9)	52.5 (4.3)	48.4 (4.0)	50.4 (3.0)	40.0 (2.3)
REST	(Insertion of 50 nt between exons 5 and 6)	ND	ND	ND	ND	ND	6.8 (1.0)
TERT	Alpha (deletion of 36 nt at the 5′ end of exon 6)	<5.0	<5.0	ND	ND	ND	19.9 (10.0)
VEGF	VEGF121 (deletion of 273 nt, including exons 6 and 7)	41.1 (4.7)	73.5 (7.1)	36.6 (7.4)	49.0 (5.1)	61.3 (8.3)	75.2 (8.9)
	VEGF206 (inclusion of exons 5 through 8)	ND	ND	ND	ND	ND	<5.0
Average		48.4	34.1	35.7	35.2	31.3	35.5

The values are for the variant described to the right of each gene. Relative percent expression of each variant was obtained from four separate reverse transcription reactions with each reverse transcription reaction PCR amplified in replicates of four ($n = 16$). The standard deviation between the replicates is in parentheses. Values above the level of detection but below the LOQ of 5.0%, are indicated as <5.0. Values >95.0% were marked with >95.0, except where only one variant was detected in which case the ASV is 100.0%. Variants not detected are marked ND. Averages for overall tissue splicing do not include VEGF data.

samples or show tissue-specific low-level expression. Of these four ASVs, three, Keratin 8, NME1 and PCBP4, have only been recently identified and lack a prior tissue-specific distribution profile. Keratin 8, recently discovered in non-small cell lung cancer (29), had an average ASV expression level of 7.1% and was consistently detected in all samples, with the exception of the brain, which expressed the full-length KRT8 exclusively. NME1 (non-metastatic cells 1) ASV was identified in two EST screening papers, indicating its potential use as a cancer marker (28,30). This variant was not detected in any normal tissues. PCBP4, a poly(rC)-binding protein gene, ASV was detected in all tissues with the highest expression in the brain and heart samples at 39.2 and 36.2%, respectively, the remaining samples had expression levels between 21.8 and 10.6%. Recently, conversion of PCBP4 to this deletion variant, PCBP4a, has been correlated with lung cancer (31). The fourth minor ASV, REST, a RE1-silencing transcription factor, was expressed exclusively in the testis at 6.8%. Previously, this ASV had been detected in lung cancer cell lines expressed at similar levels as the wild-type variant as well as in brain at such low abundance that it was only detected after using nested PCR (32).

The final gene analyzed was ACTN-4, in which the alternatively spliced 8′ exon was detected in all tissue samples with two samples showing a >50% expression, the testis and brain, at 51.2 and 71.1%, respectively. This variant, similar to those with low-level expression, has only recently been identified in lung cancer cell lines, where the ASV containing the 8′ exon was only detected in small cell lung cancer, with trace amounts in testis and the brain (33). Since this ASV was detected in all tissue samples we tested, in contrast to the previously reported expression profile data, additional analysis was performed to confirm the expression level of the 8′ exon in ACTN-4.

Paired extension primer correlation

To confirm the expression levels of the 8′ exon, ACTN-4 was assayed using two extension primers, one coming from the 5′ side of exon 8′, as originally used, and one from as the 3′ side (Figure 2). Both primers were designed adjacent to unique bases in the 8′ exon. The samples were multiplexed, a process that allows the two extension primers to be assayed simultaneously in the same reaction reducing procedural variations and increasing throughput. The extension primers of the multiplexed assays were designed with a different termination mixture so that the mass spectra for both the forward and reverse product of each variant would not interfere with each other. When both forward and reverse extension product values of the 8′ ACTN-4 exon were compared with each other (Figure 3) both values correlated well. With the values verified using bidirectional analysis, we suspected that the disparity of our results with previously published data was a result of the formation of heteroduplexes during PCR.

ACTN-4 exon 8′ heteroduplex analysis

The mutually exclusive splice variant of ACTN-4 is inherently difficult to assay via conventional PCR and gel analysis techniques. ASVs are generally analyzed by comparing the migration of the variants, which is the result of their molecular weight difference by either a deletion or insertion. For ACTN-4, the mutually exclusive exon is 83 nt, the same length as the constitutive exon 8 precluding it from this type of analysis. Using a restriction enzyme that digests one variant but not the other can create a difference in migration between the two variants, as was performed with ACTN-4 (33). Confounding problems of this adaptation include heteroduplexes. When two sequences have a high percent of homology, they are prone to form heteroduplexes during PCR. Heteroduplexes are not digested fully as only one strand of the DNA has the correct restriction site. Exons 8 and 8′ are 74.4% identical, while the entire amplicon between the forward and reverse primer (223 nt) has only 21 different residues resulting in a >90% identity.

To determine whether heteroduplexes were formed during conventional analysis, the PCR product using the same PCR primers as the mass extend assay was assayed on a non-denaturing 10% polyacrylamide gel. This type of heteroduplex

Figure 2. Forward and reverse extension primers for ACTN-4 variants. Extension primer 1 on the 5′ side of the mutually exclusive exon is the forward primer with CA indicating the extended bases for exon 8 and T indicating the extended base for exon 8′. Extension primer 2 on the 3′ side of the mutually exclusive exon is the reverse primer with CG indicating the extended bases for exon 8 and A indicating the extended base for exon 8′. All reactions were multiplexed in the same well and assayed together. An AGT termination mixture (ddATP, dCTP, ddGTP and ddTTP) was used for the primer extension reaction.

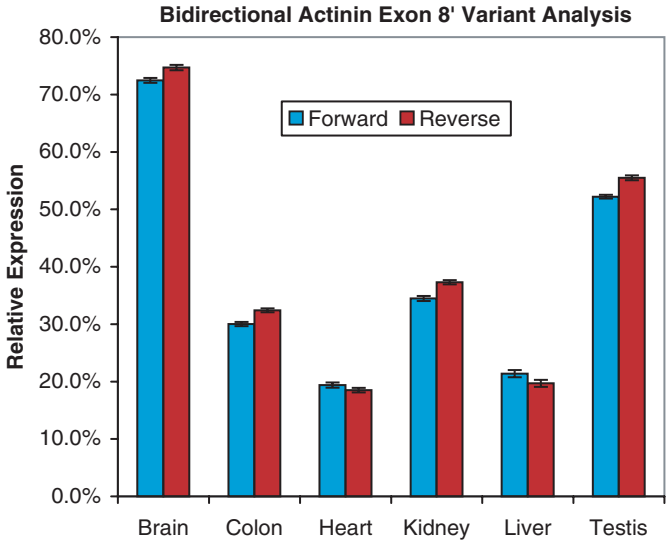

Figure 3. Bidirectional ACTN-4 exon 8′ variant analysis. Values for the forward and reverse primers were obtained from the 5′ extension primer and the 3′ extension primer, respectively. Error bars show the standard deviation.

mobility analysis (HMA) (34) allows separation of duplexes that would ordinarily appear as a single band on an agarose gel, as was seen for these samples (data not shown). HMA on two samples in lanes 4 and 5 produced three bands of different migration with one major band, 3, and two minor bands, 1 and 2 (Figure 4a). To determine whether these additional bands were heteroduplexes, aliquots of the samples run in lanes 4 and 5 were diluted 1:20 and subjected to a single PCR cycle, a process that allows heteroduplexes to denature so that correct duplexes can be re-established (35). As expected, the two-heteroduplex bands were absent in lanes 2 and 3. Finally, to confirm the heteroduplex analysis results and to ascertain the exact duplexes formed, the three bands from lanes 4 and 5 were extracted from the acrylamide and sequenced from both directions. The sequencing results confirmed that bands 1 and 2 were heteroduplexes formed by exon 8′ forward/8 reverse and exon 8 forward/8′ reverse, respectively (Figure 4b). Band 3 was a homoduplex of both exon 8 forward/8 reverse. Band 3 may also contain a homoduplex of exon 8′ forward/8′ reverse at a much smaller amount compared with exon 8 forward/8 reverse so that it is not directly detectable by sequencing.

VEGF variant analysis

After demonstrating the ability of this method to detect simple slice variants, we then moved to a more complex example, VEGF, a gene that plays a crucial role in angiogenesis, the formation of new capillaries (36). With such a vital role in vascularization and its correlation to cell proliferation, VEGF has long been considered an important factor in tumorigenesis with a role in every major type of cancer (37) and thus a natural target for therapy. VEGF has six splice variants (Figure 5a) ranging from 121 to 206 amino acids. Because of the complexity of the splicing, a single extension primer is not sufficient to identify all variants; as a result two extension primers were used, each one giving a portion of the data that together would give the relative abundance of each variant to all the variants (Figure 5b). Extension primer 1 was used to determine the relative expression of 121 and 145 relative to the sum of VEGF 165, 183, 189 and 206. The second extension primer was designed to detect only the variants not distinguished in the first extension reaction (165, 183, 189 and 206). This allowed us to take the individual results for VEGF 165, 183, 189 and 206 and back calculate the percent of each variant relative to all the variants (Figure 5c).

The data obtained from the two extension primers correlated with previous tissue profile data, with 121 and 165 representing the bulk of the variants in all tissues (38). Historically, as the minor variants were being revealed, two variants, 183 and 189, were initially not differentiated because of their small size difference, 18 nt. Lei *et al*. (39) had suggested that previously identified 189 in the heart was probably 183. Both of the variants 183 and 189 were identified to be most highly expressed in the heart at 11.2 and 8.6%, respectively, clarifying that both 183 and 189 are present in the heart. The remaining variants, 145 and 206, are rarely expressed and ostensibly limited to placental cells (40); however, there was a detectable level of 206 in testis.

DISCUSSION

We have shown that primer extension coupled with MALDI-TOF MS can be applied to high-throughput ASV detection and quantification. The tissue profiles demonstrated by this method are not meant to be definitive values, rather an illustration of the power of this method to generate statistically significant

Figure 4. Heteroduplex mobility analysis and sequencing of ACTN-4. (**a**) 10% TBE non-denaturing acrylamide gel analysis. Lane 1, 100 bp marker; lane 2, 1:20 dilution of lane 4 sample with 1 cycle of PCR; lane 3, 1:20 dilution of lane 5 sample with 1 cycle of PCR; lanes 4 and 5, undiluted PCR amplification product using ACTN-4 primers from Table 1. Lanes 2 and 3 have one band at 223 nt (band 3); lanes 4 and 5 have the same band at 223 nt and two additional bands with less mobility, bands 1 and 2. These bands were extracted from the acrylamide gel using electro-elution and sequenced using the forward and reverse primers used in the PCR amplification. (**b**) Sequencing analysis shows the unique sequences to exon 8 in green and exon 8′ in red.

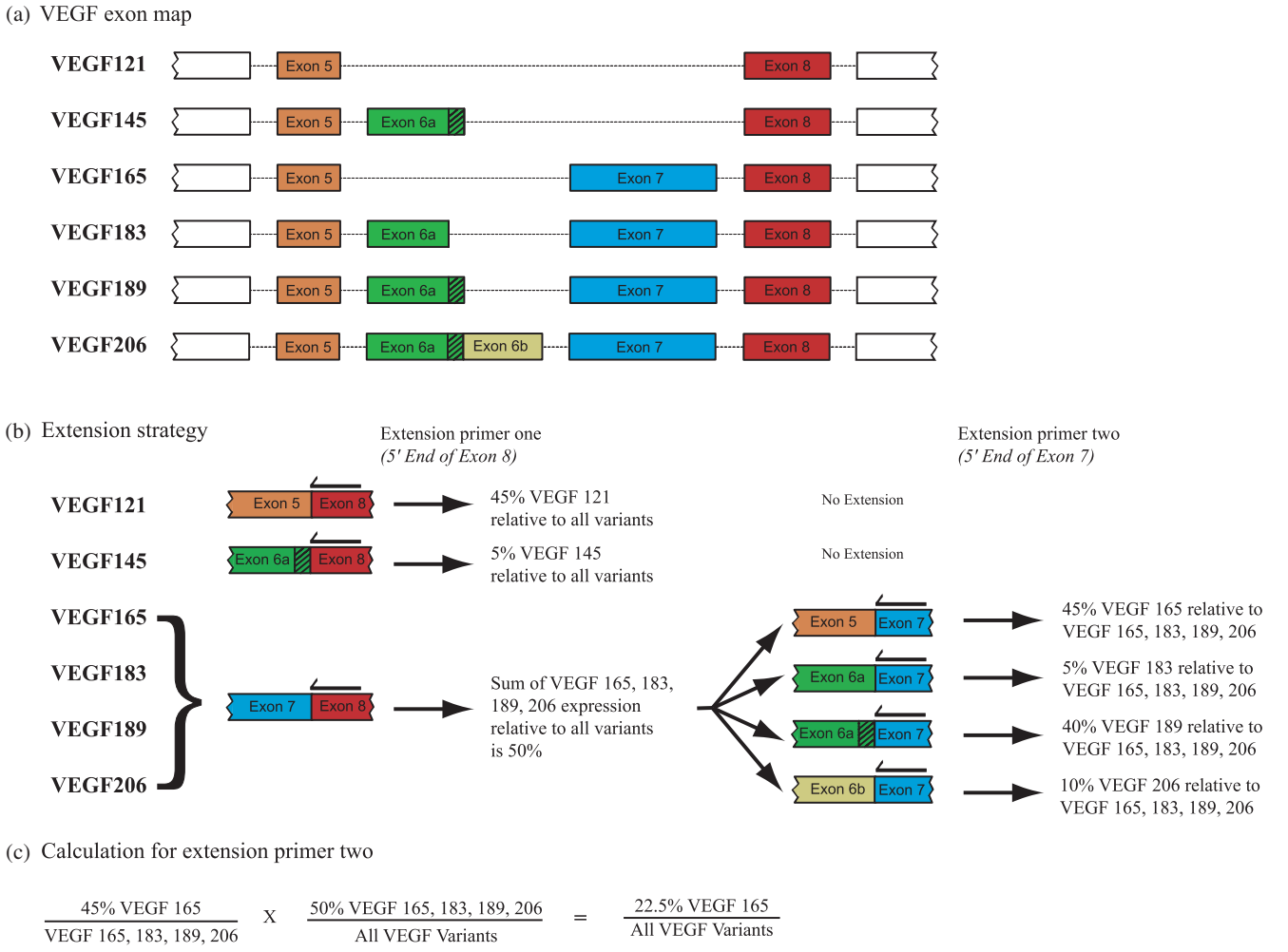

Figure 5. VEGF exon map. (**a**) Different ASVs. (**b**) Two extension primers designed to distinguish all six ASVs. Primer one targeting the 5′ end of exon eight was used to distinguish VEGF121, VEGF145 and the sum of VEGF165, VEGF183, VEGF189 and VEGF206. Primer two targeting the 5′ end of exon seven was used to distinguish VEGF165, VEGF183, VEGF189 and VEGF206 (primer 2 does not detect VEGF121 and VEGF14). (**c**) Calculation of each VEGF ASVs. Data used in this figure are for illustration only and do not reflect data from any sample assayed.

amounts of data on multiple genes, including complex splicing (VEGF). Furthermore, this method can be used to quantify ASVs with similar length and sequence that may be erroneously assayed by traditional RT–PCR gel electrophoresis method (ACTN-4). The promise of this method, coupled with the versatility of the MALDI-TOF MS platform, indicates a potential for information-rich analysis for a multitude of uses. Use of this method, however, is limited to those

ASVs that have been characterized at the sequence level and is currently not able to identify new variants. Thus, this method can function as a complement to the enormous dataset generated from cDNA sequencing and computational analyses.

ACKNOWLEDGEMENTS

The authors thank Shengnan Jin for providing technical help. This work was supported by a grant from SEQUENOM to Boston University. Funding to pay the Open Access publication charges for this article was provided by Boston University.

Conflict of interest statement. Charles R. Cantor holds stock in SEQUENOM, the makers of MassARRAY products. He is an executive officer and director of the company, and is currently conducting research sponsored by this company. The other authors have no conflict to declare.

REFERENCES

1. Berget,S.M., Moore,C. and Sharp,P.A. (1977) Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proc. Natl Acad. Sci. USA*, **74**, 3171–3175.
2. Sharp,P.A. (1994) Split genes and RNA splicing. *Cell*, **77**, 805–815.
3. Mironov,A.A., Fickett,J.W. and Gelfand,M.S. (1999) Frequent alternative splicing of human genes. *Genome Res.*, **9**, 1288–1293.
4. Johnson,J.M., Castle,J., Garrett-Engele,P., Kan,Z., Loerch,P.M., Armour,C.D., Santos,R., Schadt,E.E., Stoughton,R. and Shoemaker,D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141–2144.
5. Modrek,B. and Lee,C. (2002) A genomic view of alternative splicing. *Nature Genet.*, **30**, 13–19.
6. Strehler,E.E. and Zacharias,D.A. (2001) Role of alternative splicing in generating isoform diversity among plasma membrane calcium pumps. *Physiol. Rev.*, **81**, 21–50.
7. Muro,A.F., Iaconcig,A. and Baralle,F.E. (1998) Regulation of the fibronectin EDA exon alternative splicing. Cooperative role of the exonic enhancer element and the 5′ splicing site. *FEBS Lett.*, **437**, 137–141.
8. Lou,H. and Gagel,R.F. (1998) Alternative RNA processing—its role in regulating expression of calcitonin/calcitonin gene-related peptide. *J. Endocrinol.*, **156**, 401–405.
9. Hwu,W.L., Yeh,H.Y., Fang,S.W., Chiang,H.S., Chiou,Y.W. and Lee,Y.M. (2003) Regulation of GTP cyclohydrolase I by alternative splicing in mononuclear cells. *Biochem. Biophys. Res. Commun.*, **306**, 937–942.
10. Garcia-Blanco,M.A., Baraniak,A.P. and Lasda,E.L. (2004) Alternative splicing in disease and therapy. *Nat. Biotechnol.*, **22**, 535–546.
11. Faustino,N.A. and Cooper,T.A. (2003) Pre-mRNA splicing and human disease. *Genes Dev.*, **17**, 419–437.
12. Xu,Q. and Lee,C. (2003) Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences. *Nucleic Acids Res.*, **31**, 5635–5643.
13. Xu,Q., Modrek,B. and Lee,C. (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res.*, **30**, 3754–3766.
14. Huang,H.D., Horng,J.T., Lee,C.C. and Liu,B.J. (2003) ProSplicer: a database of putative alternative splicing information derived from protein, mRNA and expressed sequence tag sequence data. *Genome Biol.*, **4**, R29.
15. Hui,L., Zhang,X., Wu,X., Lin,Z., Wang,Q., Li,Y. and Hu,G. (2004) Identification of alternatively spliced mRNA variants related to cancers by genome-wide ESTs alignment. *Oncogene*, **23**, 3013–3023.
16. Yeo,G., Holste,D., Kreiman,G. and Burge,C.B. (2004) Variation in alternative splicing across human tissues. *Genome Biol.*, **5**, R74.
17. Zhu,J., Shendure,J., Mitra,R.D. and Church,G.M. (2003) Single molecule profiling of alternative pre-mRNA splicing. *Science*, **301**, 836–838.
18. Ross,P., Hall,L. and Haff,L.A. (2000) Quantitative approach to single-nucleotide polymorphism analysis using MALDI-TOF mass spectrometry. *BioTechniques*, **29**, 620–626, 628–629.
19. Ulaner,G.A., Hu,J.F., Vu,T.H., Giudice,L.C. and Hoffman,A.R. (1998) Telomerase activity in human development is regulated by human telomerase reverse transcriptase (hTERT) transcription and by alternate splicing of hTERT transcripts. *Cancer Res.*, **58**, 4168–4172.
20. Cerezo,A., Kalthoff,H., Schuermann,M., Schafer,B. and Boukamp,P. (2002) Dual regulation of telomerase activity through c-myc-dependent inhibition and alternative splicing of hTERT. *J. Cell Sci.*, **115**, 1305–1312.
21. Wright,W.E., Piatyszek,M.A., Rainey,W.E., Byrd,W. and Shay,J.W. (1996) Telomerase activity in human germline and embryonic tissues and cells. *Dev. Genet.*, **18**, 173–179.
22. Ulaner,G.A., Hu,J.F., Vu,T.H., Oruganti,H., Giudice,L.C. and Hoffman,A.R. (2000) Regulation of telomerase by alternate splicing of human telomerase reverse transcriptase (hTERT) in normal and neoplastic ovary, endometrium and myometrium. *Int. J. Cancer*, **85**, 330–335.
23. Reeve,J.G., Xiong,J., Morgan,J. and Bleehen,N.M. (1996) Expression of apoptosis-regulatory genes in lung tumour cell lines: relationship to p53 expression and relevance to acquired drug resistance. *Br. J. Cancer*, **73**, 1193–1200.
24. Tu,Y., Renner,S., Xu,F., Fleishman,A., Taylor,J., Weisz,J., Vescio,R., Rettiig,M., Berenson,J., Krajewski,S. *et al.* (1998) BCL-X expression in multiple myeloma: possible indicator of chemoresistance. *Cancer Res.*, **58**, 256–262.
25. Nagane,M., Levitzki,A., Gazit,A., Cavenee,W.K. and Huang,H.J. (1998) Drug resistance of human glioblastoma cells conferred by a tumor-specific mutant epidermal growth factor receptor through modulation of bcl-XL and caspase-3-like proteases. *Proc. Natl Acad. Sci. USA*, **95**, 5724–5729.
26. Olopade,O.I., Adeyanju,M.O., Safa,A.R., Hagos,F., Mick,R., Thompson,C.B. and Recant,W.M. (1997) Overexpression of BCL-x protein in primary breast cancer is associated with high tumor grade and nodal metastases. *Cancer J. Sci. Am.*, **3**, 230–237.
27. Gonzalez-Garcia,M., Perez-Ballestero,R., Ding,L., Duan,L., Boise,L.H., Thompson,C.B. and Nunez,G. (1994) Bcl-XL is the major bcl-x mRNA form expressed during murine development and its product localizes to mitochondria. *Development*, **120**, 3033–3042.
28. Wang,Z., Lo,H.S., Yang,H., Gere,S., Hu,Y., Buetow,K.H. and Lee,M.P. (2003) Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer. *Cancer Res.*, **63**, 655–657.
29. Tojo,Y., Bandoh,S., Fujita,J., Kubo,A., Ishii,T., Fukunaga,Y., Ueda,Y., Yang,Y., Wu,F., Huang,C.L. *et al.* (2003) Aberrant messenger RNA splicing of the cytokeratin 8 in lung cancer. *Lung Cancer*, **42**, 153–161.
30. Brentani,H., Caballero,O.L., Camargo,A.A., da Silva,A.M., da Silva,W.A.,Jr, Dias Neto,E., Grivet,M., Gruber,A., Guimaraes,P.E., Hide,W. *et al.* (2003) The generation and utilization of a cancer-oriented representation of the human transcriptome by using expressed sequence tags. *Proc. Natl Acad. Sci. USA*, **100**, 13418–13423.
31. Pio,R., Zudaire,I., Pino,I., Castano,Z., Zabalegui,N., Vincent,S., Garcia-Amigot,F., Odero,M.D., Lozano,M.D., Garcia-Foncillas,J. *et al.* (2004) Alpha CP-4, encoded by a putative tumor suppressor gene at 3p21, but not its alternative splice variant alpha CP-4a, is underexpressed in lung cancer. *Cancer Res.*, **64**, 4171–4179.
32. Coulson,J.M., Edgson,J.L., Woll,P.J. and Quinn,J.P. (2000) A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. *Cancer Res.*, **60**, 1840–1844.
33. Honda,K., Yamada,T., Seike,M., Hayashida,Y., Idogawa,M., Kondo,T., Ino,Y. and Hirohashi,S. (2004) Alternative splice variant of actinin-4 in small cell lung cancer. *Oncogene*, **23**, 5257–5262.
34. Upchurch,D.A., Shankarappa,R. and Mullins,J.I. (2000) Position and degree of mismatches and the mobility of DNA heteroduplexes. *Nucleic Acids Res.*, **28**, e69.
35. Becker-Andre,M. and Hahlbrock,K. (1989) Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcript titration assay (PATTY). *Nucleic Acids Res.*, **17**, 9437–9446.

36. Ribatti,D. (2005) The crucial role of vascular permeability factor/vascular endothelial growth factor in angiogenesis: a historical review. *Br. J. Haematol.*, **128**, 303–309.
37. Hicklin,D.J. and Ellis,L.M. (2005) Role of the vascular endothelial growth factor pathway in tumor growth and angiogenesis. *J. Clin. Oncol.*, **23**, 1011–1027.
38. Robinson,C.J. and Stringer,S.E. (2001) The splice variants of vascular endothelial growth factor (VEGF) and their receptors. *J. Cell Sci.*, **114**, 853–865.
39. Lei,J., Jiang,A. and Pei,D. (1998) Identification and characterization of a new splicing variant of vascular endothelial growth factor: VEGF183. *Biochim. Biophys. Acta*, **1443**, 400–406.
40. Anthony,F.W., Wheeler,T., Elcock,C.L., Pickett,M. and Thomas,E.J. (1994) Short report: identification of a specific pattern of vascular endothelial growth factor mRNA expression in human placenta and cultured placental fibroblasts. *Placenta*, **15**, 557–561.